ISTANBUL UNIVERSITY JOURNAL OF THE SCHOOL OF BUSINESS

İSTANBUL ÜNİVERSİTESİ İŞLETME FAKÜLTESİ DERGİSİ

SNAP IT

www.ifdergisi.org

KASIM 2013 • CİLT 42 • SAYI 2 • ISSN 1303-1732 • NOVEMBER 2013 • VOLUME 42 • NUMBER 2

İSTANBUL ÜNİVERSİTESİ
İŞLETME FAKÜLTESİ
DERGİSİ

İstanbul Üniversitesi İşletme Fakültesi Dergisi, İstanbul Üniversitesi
İşletme Fakültesi'nin uluslararası resmi hakemli bilimsel dergisidir.
Nisan ve Kasım aylarında olmak üzere yılda iki kez yayımlanır.

▫ ▫ ▫

Istanbul University Journal of the School of Business is the international
official peer-reviewed journal of Istanbul University, School of Business.
The journal is semiannually published
in April and November.

ISSN 1303 – 1732

Cilt 42 ▫ Sayı 2 ▫ Kasım 2013
Volume 42 ▫ Number 2 ▫ November 2013

---

www.ifdergisi.org

İndekslenme ve Özetlenme / Abstracting and Indexing:

- EconPapers - RePEc
- EBSCO
- Index Copernicus
- Ulrich's Periodicals Directory
- International Abstracts in Operations Research
- Arastirmax Bilimsel Yayın İndeksi
- ULAKBIM - Social Sciences Database

ISTANBUL UNIVERSITY
JOURNAL OF THE SCHOOL OF BUSINESS

Kapak Tasarım / Cover Design
Ali Yüksel

BASKI VE CİLT

Extrafikir Reklamevi
Tel: 0 212 5675163
E-posta: info@extrafikir.com
Web: www.extrafikir.com

İSTANBUL ÜNİVERSİTESİ İŞLETME FAKÜLTESİ DERGİSİ

ISTANBUL UNIVERSITY JOURNAL OF THE SCHOOL OF BUSINESS

İmtiyaz Sahibi/Licensee: Erdal Tekarslan
Sorumlu Yazı İşleri Müdürü/Responsible Desk Editor: Abdullah Okumuş

Editör | Editor-in-Chief
EYÜP ÇETİN
İstanbul Üniversitesi İşletme Fakültesi
Istanbul University School of Business
E-mail: eycetin@istanbul.edu.tr

Yönetici Editör | Executive Editor
TİMUR KESKİNTÜRK
İstanbul Üniversitesi İşletme Fakültesi
Istanbul University School of Business
E-mail: tkturk@istanbul.edu.tr

ÜRETİM EDİTÖRLERİ
PRODUCTION EDITORS

Hilal Özen
İstanbul Üniversitesi İşletme Fakültesi
Istanbul University School of Business
E-mail: hilaloz@istanbul.edu.tr

Ayşegül Karaeminoğulları
İstanbul Üniversitesi İşletme Fakültesi
Istanbul University School of Business
E-mail: aysegulk@istanbul.edu.tr

Selen Öztürk
İstanbul Üniversitesi İşletme Fakültesi
Istanbul University School of Business
E-mail: selenoz@istanbul.edu.tr

DANIŞMA KURULU
ADVISORY BOARD

Hamparsum Bozdogan
University of Tennessee

John L. Casti
International Institute for Applied Systems Analysis

Arjun K. Gupta
Bowling Green State University

C. Warren Neel
University of Tennessee

Thomas L. Saaty
University of Pittsburgh

Wayne L. Winston
Indiana University

EDİTÖRLER KURULU / EDITORIAL BOARD

Ahmet Cevat Acar, Istanbul University
İbrahim Zeki Akyurt, Istanbul University
Gültekin Altuntaş, Istanbul University
C. Cüneyt Arslantaş, Istanbul University
Muhteşem Baran, Istanbul University
John Chen, Bowling Green State University
Seungjin Choi, Pohang University Sci.&Tech.
Biagio Ciao, University of Milan-Bicocca
Nalan Cinemre, Mimar Sinan Fine Arts University
Tunçhan Cura, Istanbul University
Adele Diederich, Jacobs University Bremen
Öner Esen, Istanbul University
Özgür Esen, Istanbul University
Orhan Göker, Istanbul University
İbrahim Güngör, Akdeniz University
Talha Harcar, Penn State University
Hede Helfrich-Hölter, University of Hildesheim
J. Andrew Howe, TransAtlantic Petroleum
Lerzan Kavut, Istanbul University
İsmail Kaya, Istanbul University
Erdener Kaynak, Pennsylvania State University at Harrisburg
Sema Kurtuluş, Istanbul University
Fatma Lorcu, Trakya University
Sona Mardikyan, Boğaziçi University
Saita Massimo, University of Milan-Bicocca
Abdullah Okumuş, Istanbul University
Neyran Orhunbilge, Istanbul University
Necdet Özçakar, Istanbul University
Oya Özçelik, Istanbul University
Erhan Özdemir, Istanbul University
Hasan Pirkul, University of Texas at Dallas
Guoqi Qian, The University of Melbourne
Serra Eren Sarıoğlu, Istanbul University
Fulya Sarvan, Akdeniz University
Belkıs Seval, Istanbul University
Kazuo Shigemasu, The University of Tokyo
Karan P. Singh, University of North Texas
Halil Mete Soner, Swiss Finance Institute
Muni S. Srivastava, University of Toronto
Ahmet Şekerkaya, Istanbul University
Erdal Tekarslan, Istanbul University
Angela Testi, University of Genova
Mehpare Timor, Istanbul University
Cavide Uyargil, Istanbul University
Hayri Ülgen, Istanbul University
Vincenzo Esposito Vinzi, ESSEC Business School of Paris and Singapore
Nilgün Çil Yavuz, Istanbul University
M. Bayram Yıldırım, Wichita State University
Yunbin Zhao, University of Birmingham

İSTANBUL ÜNİVERSİTESİ İŞLETME FAKÜLTESİ DERGİSİ

ISTANBUL UNIVERSITY JOURNAL OF THE SCHOOL OF BUSINESS

İÇİNDEKİLER / CONTENTS

Cilt 42 • Sayı 2 • Kasım 2013
Volume 42 • Number 2 • November 2013

MAKALELER / ARTICLES

İstanbul Üniversitesi İşletme Fakültesi Dergisi
Istanbul University Journal of the School of Business
Cilt/Vol:42, Sayı/No:2, 2013, 219-234
ISSN: 1303-1732 – www.ifdergisi.org © 2013

Measuring domestic digital divide by using latent class analysis: A case study of Turkey

Çiğdem Arıcıgil Çilan[1]
Department of Quantitative Methods
School of Business
Istanbul University

Muhlis Özdemir[2]
Department of Quantitative Methods
School of Business
Istanbul University

Abstract

With the rapid change and development of information technologies, use of technology and access levels to technology have now gained importance in determining the development levels of countries. Countries with access and ability to use information technologies are considered developed. The differences in use of technology between countries and/or differences in access to and use of information technology between individuals in a country have introduced the "digital divide" concept. The difference between countries is called "international digital divide" whereas the differences between individuals in a country are called "domestic digital divide". Based on variables that determine the levels of access to and use of information technologies in Turkey, the purpose of this study is to analyze the country's "domestic digital divide". For this purpose, data about Turkey in European Social Survey were put to a Latent Class Analysis used in Categorical Data Analysis.

Keywords: *Digital Divide, Information Technologies, Development, Latent Class Analysis, Categorical Data Analysis*

Gizli sınıf analizi ile sayısal uçurumun ölçümlenmesi: Türkiye örneği

Özet

Bilişim teknolojilerinde yaşanan hızlı değişim ve gelişimle birlikte artık ülkelerin gelişmişlik düzeylerinin belirlenmesinde teknoloji kullanım oranları ve bu teknolojilere erişim düzeyleri önem kazanmaktadır. Ülkeler bilgi teknolojilerine erişebildikleri ve bu teknolojileri kullanabildikleri ölçüde gelişmiş sayılmaktadır. Ülkeler arasındaki teknoloji kullanımı farklılıkları ve/veya bir ülkenin bireyler arasındaki bilgi teknolojilerinin kullanımı ve erişimindeki farklılıklar "sayısal uçurum" kavramını ortaya çıkarmıştır. Ülkeler arasındaki farklılıklar "uluslararası sayısal uçurum" ve ülkenin bireyleri arasındaki farklılıklar ise "ülke içindeki sayısal uçurum" olarak adlandırılmaktadır. Bu çalışmanın amacı Türkiye'ye ilişkin bilişim teknolojileri kullanım ve erişim düzeylerini belirleyen değişkenler temel alınarak ülke içindeki sayısal uçurumu analiz etmektir. Bu amaçla European Social Survey'den elde edilen Türkiye verilerine kategorik veri analizinde kullanılan gizli sınıf analizi uygulanacaktır.

Anahtar Sözcükler: *Sayısal Uçurum, Bilişim Teknolojileri, Gelişmişlik, Gizli Sınıf Analizi, Kategorik Veri Analizi*

1. Introduction

The rapid development and change in information technologies in Turkey has led to an increase in IT usage rates in recent years. A look into the recent increase in rates of

[1] *ccilan@istanbul.edu.tr (Ç.A. Cilan)*

[2] *muhlisozdemir@istanbul.edu.tr (M. Özdemir)*

computer and internet use reveals a geometric growth. Numerous studies emphasize that the use of information technologies in a country is parallel to its level of development. Today, the level of information technologies is considered one of the key indicators to measure development. However, the level of IT access and use vary among individuals in many countries. These differences usually stem from individuals' demographic variation. This situation known as digital divide is a result of differences in age, gender, education, income and place of residence (rural or urban). This study aims to measure the digital divide which is believed to be an outcome of such differences. For this purpose, the relevant data from the European Social Survey, conducted in Turkey in 2008 and published in 2011, were used. The number of different classes that Turkish individuals are broken up into regarding the use and access of information technology according to this data which has been put through Latent Class Analysis will be studied.

For the data from Turkish Statistical Institute's IT Usage of Households Survey were unreachable, the study uses data collected in Turkey and published on the web by the European Social Survey, which is an internationally accepted data collection centre. There are a total of 5 variables in the survey related to information technologies: "weekly television watching hours", "internet usage rates for personal purposes," "landline ownership", "mobile phone ownership", "phone call via the internet". Latent Class Analysis was applied based these 5 variables.

This study will try to explain the concept of digital divide by using the Latent Class Analysis technique. The following parts are literature review, latent class analysis, hypotheses of latent class analysis, conditional independence model in latent class analysis, goodness of fit, restricted latent class models, research, and results.

2. Literature Review

With the rapid developments in information technologies the digital divide has become an important issue. To date, studies on digital divide have been mainly based on the digital divide between developed, developing and underdeveloped countries while there are also studies are based on differences between various geographical regions. As countries become more advanced, the information gap will also decrease. There is a high correlation between the income per capita and the usage of information technology [1].

The term digital divide was used for the first time in the 1990s by Larry Irving, the former secretary general of National Telecommunications Infrastructure Administration [2]. Initially, the digital divide was defined based on computer ownership or knowing how to use the computer [3]. Recently, it has been used as a new form of computer and internet information technologies. Many researchers define digital divide as the obstacles encountered in accessing information technology systems [4]. According to the OECD definition, digital divide describes inequality experienced by different socioeconomic level of individuals, companies or countries in accessing and using information technology [5]. Researchers and various theories have sought to develop methodologies to study digital divide. Corrocher and Ordanini's [6] 2002 study emphasizes that Ricci's research sets the preliminary theoretical framework. Previously discussed only by developed countries, the concept of digital divide is today considered important in many countries. International platforms such as the United Nations, World Bank and OECD discuss the different rates of access to information resulting from the use of information technologies [7].

The studies that aim to measure and define the digital divide show some differences arising from the researchers. Studies on digital divide tend to divide in two groups. One group measures digital divide with statistical methods while the other group focuses on the development of digital divide and the definition of variables used to identify the digital divide [8]. Using micro data sets to identify the factors, some studies determined that digital divide is not a result of individual characteristics but a person's living space (rural or urban) cause digital divide [9]. A look into previous studies show that the most widely used statistical methods in measuring digital divide are Regression and Correlation Analysis, Factor Analysis, Multivariate Statistical Methods, Multidimensional Scaling, Canonical Correlation Analysis, Discriminant Analysis and Manova. However, the number of local researches is low while cross-country comparisons are numerous. The reason is that it is very difficult or even impossible to reach local data. Because of this constraint, some researchers try to explain the digital divide by using single-variable data sets in Internet access rates or rates of use of a particular website. The most frequent variables used in descriptive studies are age, gender, education level, income level, geographic location, internet access, computer ownership and mobile phone ownership. The most important variables used in studies on international differences are teledensity, the number of PCs, the number of web sites, the number of hosts, the number of internet users, bandwidth, user language, languages of websites, labour force employed in information technologies and the share of information communication technologies in exports [7]. Education level is identified as one of the most important factors in digital divide. Empirical studies show that younger generation adapts more easily to information technologies and assert that easy adaptation is related to their different behaviours, attitudes and skills. Education of the younger generation enables them to follow technology closely [10]. A research by Forman et al [11] indicate that population in the place of residence, population density and the urban-rural divide lead to digital divide and that internet access is faster in urban areas than in rural areas for the technological infrastructure is better in cities. Given the costs of infrastructure needed in rural areas, access can be said to be cheaper in rural areas.

A look into studies on the use of information technologies in Turkey shows that the first study was conducted in 1997 by Turkish National Information Infrastructure Project under the title "Assessment of Use and Skills"[12]. This research was carried out by Turkish Statistical Institute[13]. It states that studies on domestic digital divide were conducted at the end of 1990s and the beginning of 2000s; however, the number and content of these studies are limited. Information Technology Penetration and Usage Survey again in 2000 under the name of a field study was carried out. In 2000 a study titled Information Technology Penetration and Usage Survey was published. The study consisted of 600 households from 68 cities and 165 provinces in 7 geographical regions. This research was again carried out by the Turkish Statistics Institute [13]. Özcivelek et al [14] in 2000 examine digital divide in the world and in Turkey. Oruç and Arslan [15] in 2002 developed a strategic plan for the elimination of digital divide. In 2003, Aşkar analysed digital divide in education [16]. Turkish Statistical Institute has conducted "Use of IT in the Household Survey" since 2005. Uçkan [17] in 2008 published "Digital Divide and Knowledge Gap". Seferoğlu et al [18] discuss in their 2008 study situation in Turkey and possible policies to fight against the digital divide in Turkey. The State Planning Organization, today known as Ministry of Development of Turkey, published a report titled "Information Society Strategy" covering the 2006-2010 period. In 2010 Turkish Statistical Institute carried out a research titled "Use of IT in Businesses Research" to compare the rates of internet use in Turkish businesses.

According to the 2010 Household IT Usage Survey conducted in line with EU norms by the Turkish Statistical Institute with the contribution of European Union Statistics Office, computer and Internet usage rates are rapidly increasing in Turkey. According to the survey, Internet usage rate is 37.6% in 2010 can be seen in Figure 1.

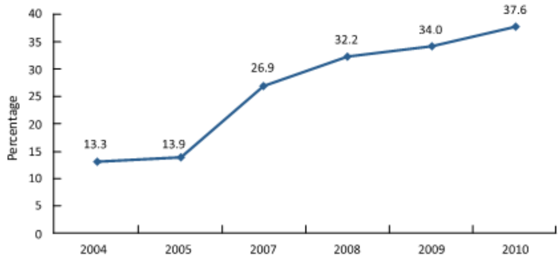

Figure 1 Internet Use in Years, 2004-2010

Source: Turkish Statistical Institute IT Usage of Households Survey 2010 [13]

As the impact of geographic region on digital divide is accepted, the survey results prove this situation right. The gap between urban and rural use rates differs noticeably. According to the survey, areas with a population under 20.000 are defined rural whereas those with a population above 20.000 are defined urban. As for this definition, 71% of the population lives in cities while 29% live in rural areas. Urban and rural use rates can be seen in Figure 2 and Figure 3.

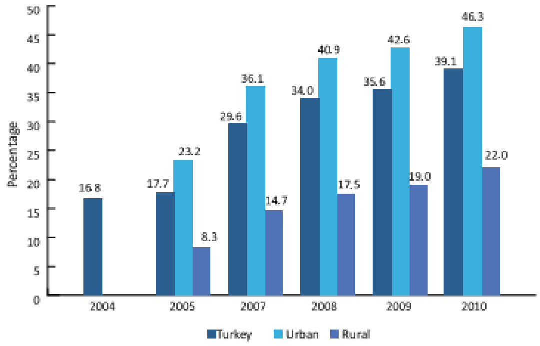

Figure 2 Rural and Urban Computer Use, 2004-2010

Source: Turkish Statistical Institute IT Usage of Households Survey 2010 [13]

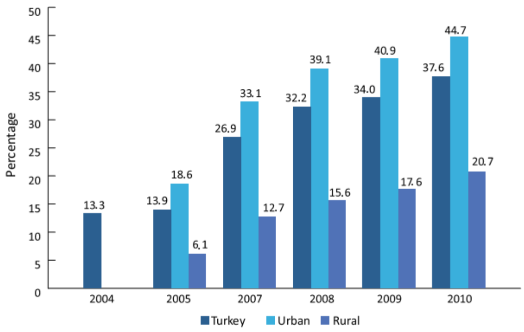

Figure 3 Rural and Urban Internet Use, 2010

Source: Turkish Statistical Institute IT Usage of Households Survey 2010 [13]

At the same time, the gender variable is effective on digital divide in Turkey. The survey again shows that rates of internet use differ according to gender and age groups. The highest rate of internet use is in 16-24 year-old age group. In addition, as for rates of internet use in all age groups, women lag behind men. This proves that there is a digital divide in Turkey between women and men. These rates can be seen in Figure 4.

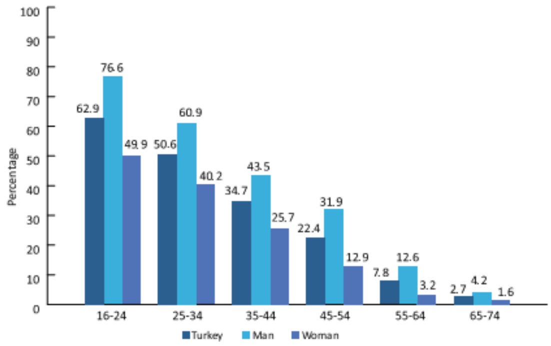

Figure 4 Internet Use in Age Groups, 2010

Source: Turkish Statistical Institute IT Usage of Households Survey 2010 [13]

Figure 5 demonstrates that rate of internet use rises together with people's educational level. It can be said that the digital divide in terms of gender has become almost non-apparent with the increase in levels of education. According to Fairlie, Beltran and Das' 2010 study [10], this improvement results from the need to use computers in education. Unfortunately, the difference is not linear as the level of education drops. The divide between men and women in internet use increases when the education level decreases.

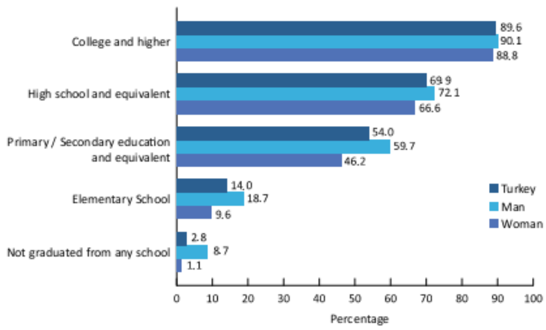

Figure 5 Internet Use in Education Levels, 2010

Source: Turkish Statistical Institute IT Usage of Households Survey 2010 [13]

3. Latent Class Analysis

In the field of Categorical Data Analysis, as an alternative to many continuous variable analyses there are methods fit to nominal and ordinal variables. Latent Class Analysis can be defined as "an alternative of Factor Analysis for categorical variables". Unlike in Factor Analysis, normal distribution assumption is not required as data are categorical.

Lazerfeld used the concept of Latent Structure Analysis to define the attitude variable in mathematical models [19,20]. Lazerfeld considers Factor Analysis as a latent structure method because it is an analysis in which latent variables are collected from continuously observed variables [19]. As categorical latent variables are derived from two or more observed variables, Latent Class Analysis is accepted as the categorical counterpart of Factor Analysis [21,22]. Latent structure methods are classified according to whether latent variable or observed variable is categorical or not can bee seen in Table 1.

Table 1 Classification of Methods with Latent and Observed Variables [23]

	Observed Variables	
Latent Variables	**Categorical**	**Continuous**
Categorical	Latent Class Analysis	Latent Profile Analysis
Continuous	Latent Feature Analysis	Factor Analysis

This study is based on Latent Class Analysis for both observed and latent variables are categorical. As Latent Class Analysis solely depends on categorical variables, it is based on the analysis of contingency tables made up of relevant categorical variables.

4. Assumptions of Latent Class Analysis

Units show Multinomial or Poisson distribution. In Poisson distribution *n* is random while in Multinominal Variable it is fixed. Samples are independent.

In the contingency tables of the categorical variables, each cell must include at least one unit. In other words, as every cell in the contingency table shows a possible result, all possible results must occur at least once. In case of a possible result that does not occur (if the cells in the contingency table show zero), a multinominal distribution cannot be guaranteed.

In this study, 5 two-category variables consist of 32 possible outcomes. This can be thought of as a questionnaire containing 5 questions with two options. A contingency table of the results of this survey will be a 5-way table with 32 $(2x2x2x2x2)$ cells. The purpose of latent class analysis here is to reduce the number of these 32 results.

5. Unrestricted Independence Model in Latent Class Analysis

A, B, C, D and E are categorical variables while i, j, k, l, m are the number of categories respectively. The restricted independence model in Latent Class Analysis is as follows (*t:* Number of categories of latent variable):

$$\pi_{ijkl}^{ABCD} = \sum_{t=1}^{T} \pi_t^X \quad \pi_{it}^{A/X} \quad \pi_{jt}^{B/X} \quad \pi_{kt}^{C/X} \quad \pi_{lt}^{D/X} \quad \pi_{mt}^{E/X} \qquad (1)$$

Here π_{ijklm}^{ABCDE} is a joint probability when π_t^X gives the possibility of every category of the (X) latent variable and these probabilities are defined as mixture proportions or latent class probabilities. If the latent variable has 2 classes $(t = 1, 2)$, latent class probabilities are shown as π_1^X *and* π_2^X are the accrual probabilities in the latent variable. There are two important factors affecting the mixture proportions: the first is the number of classes in the latent variable (t) and the second one is their relative sizes. The sizes of the latent class probabilities tell whether population is similarly dispersed among classes. The total of latent class probabilities is 1:

$$\sum_{t}^{T} \pi_t^X = 1 \qquad (2)$$

In Latent Class Analysis the numbers of estimated latent class probabilities are T-1. Latent class probabilities are also used for comparing latent class structures of two or more populations. If the compared populations' latent class probabilities differentiate, this means that the population distribution varies according to the relevant latent classes. Comparison of latent class probabilities may also mean the comparison of the latent class probabilities of the same population at different points in time. In this case, the latent probabilities obtained may provide data about how the latent probability distribution changes over time.

Another parameter in Latent Class Analysis is conditional probabilities. In model (1), the conditional probabilities are shown as follows:

$$\pi_{it}^{A/X} \quad \pi_{jt}^{B/X}\pi_{kt}^{C/X}, \ \pi_{lt}^{D/X} \ and \ \pi_{mt}^{E/X} \qquad (3)$$

Conditional probabilities are similar to factor loadings in Factor Analysis. Conditional probabilities indicate the probability of a *t* class unit to maintain a certain category in the observed variable. For each latent class, the total of number of categories of categorical variables (e.g. when working with A, B, C, D and E categorical variables) are calculated as (I+J+K+L+M) conditional probabilities. For a certain latent class, the total of conditional probabilities of each categorical variable is 1. This feature is shown as follows:

$$\sum_i \pi_{it}^{A/X} = \sum_j \pi_{jt}^{B/X} = \sum_k \pi_{kt}^{C/X} = \sum_l \pi_{lt}^{D/X} = \sum_m \pi_{mt}^{E/X} = 1 \qquad (4)$$

Similarly in a latent class analysis with 5 categorical variables, for each latent class $(I-1)+(J-1)+(K-1)+(L-1)+(M-1)$ conditional probabilities should be estimated in this case, the number of parameters to estimate in an unrestricted Latent Class Analysis is; (T-1) latent class probabilities and as many as $(I-1)+(J-1)+(K-1)+(L-1)+(M-1)$ conditional probabilities for each class (T) of the latent Variable is a total of

$$(T-1)+T[[(I-1)+(J-1)+(K-1)+(L-1)+(M-1)]]$$

$$=T+T(I+J+K+L+M-5)-1]$$

$$=T(I+J+K+L-4)-1]$$

When there are as many (T) conditional probabilities $[(I-1)+(J-1)+(K-1)+(L-1)+(M-1)]$ as the number of latent classes.

The number of present information in a 5-way contingency table is$(I*J*K*L*M-1)$.

In an unrestricted Latent Class Analysis, the number of known parameters should be larger than the number of estimated parameters.

$$(I*J*K*L*M-1) > T(I+J+K+L+M-4)-1)$$

To reach the inequality above, the maximum value of the number of latent class (T) should be restricted.

Accordingly, the degree of freedom (df.) used to test the model is calculated as:

$$\text{d.f.} = (\text{IJKLM-1}) - (\text{T}(\text{I+J+K+L-4})\text{-1})$$

With the help of joint probabilities derived from latent class equation estimate, frequencies of the population can be estimated based on a 5-way contingency table made up of $A_i, B_{j,} C_k, D_l \ ve \ E_m$ categorical variables. In other words, we can calculate how many times each probable result is repeated.

$$\hat{F}_{ijklm} = N * \pi_{ijklm}^{ABCDE} \qquad (5)$$

Here N equals the total of observed frequencies in the contingency table

$$N = \sum_{i}^{I}\sum_{j}^{J}\sum_{k}^{K}\sum_{l}^{L} F_{ijklm} \tag{6}$$

Sometimes it is possible to encounter situations where some cells in contingency tables are zero. This is defined as sparseness in statistics. Sparseness can be a problem in determining the degree of freedom required to test the model and therefore in testing the model's goodness of fit [24].

In estimating Latent Class Models, Maximum Likelihood (MLH) method is widely used. The idea of estimating latent class probabilities and conditional probabilities by using MLH was first proposed by Goodman [25, 26]. According to the methods previously used in these studies [19,27] estimated parameters should be between 0 and 1. To indicate that the model estimated from a sample model, it could be rewritten as:

$$\hat{\pi}_{ijklmt}^{ABCDEX} = \hat{\pi}_{t}^{X} \hat{\pi}_{it}^{A/X} \hat{\pi}_{jt}^{B/X} \hat{\pi}_{kt}^{C/X} \hat{\pi}_{lt}^{D/X} \hat{\pi}_{mt}^{E/X} \tag{7}$$

When the above equation is applied to each t latent class and summed up, joint probabilities of the contingency table with $(I*J*K*L*M)$ cells are obtained.

$$\hat{\pi}_{ijklm}^{ABCDE} = \sum_{t} \hat{\pi}_{ijklmt}^{ABCDEX} \tag{8}$$

6. Goodness of Fit Measures

Model's goodness of fit can be measured with Chi-Squared Goodness of Fit Test (X^2), Likelihood Ratio (G^2) tests and standardized residuals. X^2 and G^2 tests are used to test H_0: The model fits the data" hypothesis. If the H_0 hypothesis is not rejected, model is accepted as "fit" and standardized residuals are examined to see if their absolute values rise above 2. Standard residuals, X^2 and G^2 are calculated as follows:

$$e_{ijklm} = \frac{F_{ijklm} - \hat{F}_{ijklm}}{\sqrt{\hat{F}_{ijklm}}} \tag{9}$$

$$X^2 = \sum_{ijklm} \frac{(F_{ijklm} - \hat{F}_{ijklm})^2}{\hat{F}_{ijklm}} \tag{10}$$

$$G^2 = 2\sum_{ijklm} F_{ijklm} \, ln\left(\frac{F_{ijklm}}{\hat{F}_{ijklm}}\right) \tag{11}$$

As is well-known, F_{ijklm} derived from the formula below:

$$\hat{F}_{ijklm} = N * \pi_{ijklm}^{ABCDE}$$

When the sample size is above 1000, AIC are BIC are among the criteria used in deciding which model is a better fit.

$$\text{AIC} = \text{G}^2 - 2\ \text{df}. \tag{12}$$

$$\text{BIC} = \text{G}^2 - \text{df}. * [\ln(\text{N})] \tag{13}$$

When these criteria are used in model comparisons, models with the smallest negative values should be preferred.

7. Restricted Latent Class Models

In latent class models, model estimations can be made by setting various restrictions on conditional and latent class probabilities. These restrictions are "equality restriction", "deterministic restriction" and "error rate equality restriction". When "equality restriction" is applied to conditional probabilities, indicators are called "parallel indicators". For example, when B and C variables are parallel (B&C), below restrictions are maintained.

$$\pi_{11}^{B,\ X} = \pi_{11}^{C,\ X} \text{ and } \pi_{12}^{B,\ X} = \pi_{12}^{C,\ X} \tag{14}$$

When deterministic restrictions that are applied to conditional probabilities are assumed to be applied to variable A, the conditional probability indicator for A $(\pi_{11}^{A,\ X})$ is called the "perfect indicator" and gives the below equality.

$$\pi_{11}^{A,\ X} = 1 \tag{15}$$

Error rate equality restriction for variable D can be shown as:

$$\pi_{21}^{D,\ X} = \pi_{12}^{D,\ X} \tag{16}$$

Likelihood-Ratio-Chi-Square Test is generally used in comparing restricted and restricted models or in comparing models among themselves. Here M_2 represents the second model and M_1 is the first model used in comparison.

$$G^2(M_2 \setminus M_1) = G^2(M_2) - G^2(M_1) \tag{17}$$

Statistics show the Chi-Square Distribution with

$$d.f.(M_2 \setminus M_1).\ d.f.(M_2 \setminus M_1) = d.f.(M_2) - d.f.(M_1) \tag{18}$$

When $G^2(M_2 \setminus M_1) < \chi^2_{d.f.(M_2 \setminus M_1)}$ inequality is reached, M_2 model is taken.

8. The Research

The aim of the research is to analyze the digital divide in Turkey by using Latent Class Analysis. For this purpose 2008 Round 4 data regarding Turkey published by European Social Survey was used. 2416 people were interviewed in total. In this survey conducted in Turkey, there are 5 questions (5 variables) about Information Technologies (IT). Together with these variables, the research employs other variables from digital divide literature. However, the questions in this survey were not fundamental as data for such variables in Turkey could not be reached. This is the most important limit in the study. Variables in the research are "personal use of internet/e-mail/www", "total tv viewing time on average weekday", "landline ownership", "personal mobile phone ownership", and "use of internet for telephone calls at home". The original data set of these variables have 8, 8, 2, 2, 2 categories respectively.

In the research, the 5th (Once a week), 6th (Several times a week) and 7th (Everyday) categories of the "personal use of internet/e-mail/www" variable were compiled into a new category called "frequent Internet users" and coded as 1. For those who do not use the internet or use it "once a week", a new category was created called "seldom or not users" and coded as 2. Similarly for the "total tv viewing time on average weekday" variable, people who watch TV for 3 or more hours a day were grouped as "Long period TV viewers" and coded as 1. Those who do not watch TV or watch less than 3 hours a day were grouped as "seldom or not viewers" and coded as 2. Thus, the data set is designed for a Latent Class Analysis with 5 variables each with 2 categories.

9. Results of the Research

Before the analysis "personal use of internet/e-mail/www variable was named as A, "total tv viewing time on average weekday" as B, "landline ownership" as C, "personal mobile phone ownership" as D and "use of internet for telephone calls at home" as E. The following categories were shown with code "1" which means "Yes": "Frequent internet users" in variable A, "Long time TV viewers" in variable B, "landline owners" in variable C, "mobile phone owners" in variable D and "users of internet for phone calls" in variable E. In the analysis results code "1" will be shown as "+". Negative responses to survey questions were coded "2" in the data set and shown as "-" in the results.

In summary the questions that consist of A, B, C, D and E variables are:

Variable A: Do you use the internet frequently (more than once a week)?

Possible results: Yes "+", No "-"

Variable B: Do you watch TV longer than 3 hours a day?

Possible results: Yes "+", No "-"

Variable C: Do you own a landline at home?

Possible results: Yes "+", No "-"

Variable D: Do you have a mobile phone?

Possible results: Yes "+", No "-"

Variable E: Do you use the internet to make phone calls at home?

Possible results: Yes "+", No "-"

The data set was analyzed with the LEM program and consisted of one latent variable. In order to decide on the number of its classes, Unrestricted Independence Models in which the latent variable had 1, 2, 3 and 4 classes are tested and these models were called M_0, M_1, M_2, M_3. The model results are shown in Table 2.

Table 2 Models Applied to the Cross- Classification Data

			Chi-Square	
Model	No: of Latent Classes	df.	Goodness of Fit (X^2) and p- values	Likelihood Ratio (G^2) and p-values
Independence (M_0)	1	26	1748.7663 (0.0000)	1092.7135 (0.0000)
Latent Class (M_1)	2	20	42.1214 (0.0027)	38.5440 (0.0076)
Latent Class (M_2)	3	14	27.7809 (0.0152)	23.9134 (0.0468)
Latent Class (M_3)	4	8	12.8881 (0.1158)	12.5784 (0.1272)

The model with 4 latent classes (M_3) was the only fit model following Chi-Square (X^2) Test, Likelihood Ratio (G^2) test and Standardized Residuals. In the 4-class model, the significance degree of the (X^2) test was (0.1158) whereas it was (0.1272) in the (G^2). These significance levels do not reject the H_0 hypothesis $(H_0: \textit{The model Fits the Data})$. Before discussing the results of the (M_3) model, various restrictions were applied. The results of the tested restricted models are in Table 3.

Table 3 Restricted Latent Class Models

Model	**p-value of** (X^2)	**p-value of** (G^2)
H_1: 4-Class Latent Class Model (M_3)	(0.1158)	(0.1272)
H_2: H_1 +B&C Parallel Indicators	(0.0000)	(0.0000)
H_3: H_1 +D&C Parallel Indicators	(0.0000)	(0.0000)
H_4: H_1 +A&E Parallel Indicators	(0.0001)	(0.0001)
H_5: H_1 +A as a Perfect Indicator for Class 2	(0.1158)	(0.1272)
H_6: H_1 +E as a Perfect Indicator for Class 4	(0.0902)	(0.1155)

The significance levels of the restricted models show that (M_3) is the best fit model. Therefore results of the (M_3) model will be analyzed. As the research consisted of 5 variables each with 2 categories, it is possible for interviewees to give 2^5 (32) possible responses. All the possible responses, estimated frequencies with (M_3) model and standardized residuals for each possible outcome are in Table 4.

Table 4 Observed Estimated frequencies and standardized Residuals of Possible Results (Cells)

A	B	C	D	E	Observed	Estimated	Std. Res.
+	+	+	+	+	97	93.793	0.331
+	+	+	+	-	58	54.059	0.536
+	+	+	-	+	6	4.608	0.648
+	+	+	-	-	5	6.261	-0.504
+	+	-	+	+	17	14.230	0.734
+	+	-	+	-	15	19.333	-0.986
+	+	-	-	+	1	1.998	-0.706
+	+	-	-	-	3	2.714	0.174
+	-	+	+	+	143	145.879	-0.238
+	-	+	+	-	68	70.981	-0.354
+	-	+	-	+	5	5.640	-0.269
+	-	+	-	-	8	7.662	0.122
+	-	-	+	+	12	17.415	-1.298
+	-	-	+	-	29	23.660	1.098
+	-	-	-	+	5	2.445	1.634
+	-	-	-	-	2	3.321	-0.725
-	+	+	+	+	27	31.931	-0.873
-	+	+	+	-	321	320.989	0.001
-	+	+	-	+	3	3.373	-0.203
-	+	+	-	-	181	174.263	0.510
-	+	-	+	+	6	6.867	-0.331
-	+	-	+	-	324	324.022	-0.001
-	+	-	-	+	3	2.052	0.661
-	+	-	-	-	127	133.508	-0.563
-	-	+	+	+	52	47.746	0.616
-	-	+	+	-	296	296.087	-0.005
-	-	+	-	+	3	3.768	-0.396
-	-	+	-	-	172	177.961	-0.447
-	-	-	+	+	10	8.021	0.699
-	-	-	+	-	257	256.987	0.001
-	-	-	-	+	2	2.235	-0.157
-	-	-	-	-	142	136.191	0.498

Table 4 reveals that the absolute standardized residuals do not exceed 2 and that the 4-class model is the best fit model in terms of standardized residuals.

The latent class probabilities derived from the 4-class (M_3) model are shown in Table 5.

Table 5 Latent Class Probabilities

Classes	π_t^x
1	0.1585
2	0.6159
3	0.1039
4	0.1217

Accordingly, the probability of individuals is 0.1585 to be in Class 1, 0.6159 in Class 2, 0.1039 in Class 3 and 0.1217 in Class 4. Responses in Variable A, B, C, D and E can be seen in Table 6.

Table 6 Distribution of Responses in Variable A, B, C, D and E

Response	Class 1	Class 2	Class 3	Class 4
Variable A				
+	0.7398	0.0000	0.7724	0.0000
-	0.2602	1.0000	0.2276	1.0000
Variable B				
+	0.4497	0.4954	0.3658	0.6829
-	0.5503	0.5046	0.6342	0.3171
Variable C				
+	0.6976	0.5648	1.0000	0.3265
-	0.3024	0.4352	0.0000	0.6735
Variable D				
+	0.8769	0.5798	1.0000	1.0000
-	0.1231	0.4202	0.0000	0.0000
Variable E				
+	0.4240	0.0101	0.8657	0.0000
-	0.5760	0.9899	0.1343	1.0000

In Class 1, the rate of a positive response "+" is ($\pi_{11}^{A/X} = 0.7398$) for "Do you use the internet frequently" (Variable A), ($\pi_{11}^{B/X} = 0.4497$) for "Do you watch TV longer than 3 hours a day" (Variable B), ($\pi_{11}^{C/X} = 0.6976$) for "Do you own a landline at home", ($\pi_{11}^{D/X} = 0.8769$) for "Do you have a mobile phone" and ($\pi_{11}^{E/X} = 0.4240$) for "Do you use the internet to make phone calls at home".

In Class 2, the rate of a positive response "+" is ($\pi_{12}^{A/X} = 0.0000$)for "Do you use the internet frequently" (Variable A), ($\pi_{12}^{B/X} = 0.4954$) for "Do you watch TV longer than 3 hours a day" (Variable B), ($\pi_{12}^{C/X} = 0.5648$)for "Do you own a landline at home", (0.5798) for "Do you have a mobile phone" and ($\pi_{12}^{E/X} = 0.0101$) for "Do you use the internet for phone calls at home".

In Class 3, the rate of a positive response "+" is ($\pi_{13}^{A/X} = 0.7724$)for "Do you use the internet frequently" (Variable A), ($\pi_{13}^{B/X} = 0.3658$) for "Do you watch TV longer than 3 hours" (Variable B), ($\pi_{13}^{C/X} = 1.0000$) for "Do you own a landline at home", ($\pi_{13}^{D/X} = 1.0000$) for "Do you have a mobile phone" and($\pi_{13}^{E/X} = 0.8657$) for "Do you use the internet to make phone calls at home".

In Class 4, the rate of a positive response "+" is ($\pi_{14}^{A/X} = 0.0000$) for "Do you use the internet frequently" (Variable A), ($\pi_{14}^{B/X} = 0.6829$) for "Do you watch TV longer than 3 hours a day" (Variable B), ($\pi_{14}^{C/X} = 0.3265$) for "Do you own a landline at home", ($\pi_{14}^{D/X} = 0.1000$) for "Do you have a mobile phone" and($\pi_{14}^{E/X} = 0.0000$) for "Do you use the internet to make phone calls at home".

Today, variables A and E are considered two of the most important variables which determine whether or not level of access to and use of Information Technology is high. Therefore, the classes with high conditional probabilities of these variables can be named "classes with high levels of access to and use of information technologies" while classes with low probabilities can be named "classes with low levels of access to and use of information technologies". In classes 1 and 3, "the probabilities to use the internet frequently (Variable A)" are (0.7398) and (0.7724) respectively. In the same classes, "the probabilities to use the internet to make phone calls (Variable E)" are (0.4240) and (0.8657) respectively. The probabilities are very close to zero in other classes. Accordingly, Class 1 and Class 3 can be called "high IT competence" classes. In terms of "mobile phone ownership (Variable C)", the probabilities for Class 1 and Class 3 are (0.8769) and (1.0000). The probability of those in Class 1 to watch TV more than 3 hours a day is (0.4497) while the same probability for Class 3 is (0.3658). The most competent internet users are in Class 3 in which the ratio of long time TV viewers is lower than in other classes. The proportion of landline owners is (1.0000) in Class 3 and (0.6976) in Class 1.

According to these results, Class 3 can be called "highest IT competence" whereas Class 1 can be called "high IT competence". However, as class probabilities reveal, the probability of individuals to be places in Class 3 and Class 1 is low $(\hat{\pi}_1^X = 0.1585 \quad \hat{\pi}_3^X = 0.1039)$.

Individuals with the "low levels of IT" are in Class 2 and Class 4. The probabilities of these individuals to "use the internet frequently (Variable A)" and to "use the internet to make phone calls at home (Variable E)" are very close to zero. The individuals in Class 4 have a higher probability with (0.6829) to watch long hours of TV than individuals in Class 2. In addition the probability to own a mobile phone is (1.0000) in Class 4, which means that everyone in this class may have a mobile one. In Class 4 the probability of landline ownership is lower (0.3265) than in Class 2(0.5648). Considering that mobile phone ownership and TV content in Turkey are very much affected by contemporary pop culture, Class 4 can be named "low level of IT competence and highly affected by pop culture" and Class 2 can be named "low level of IT competence".

10. Conclusion

According to the results of analysis, a domestic digital divide can be said to exist in Turkey. The results of the Latent Class Analysis show that Turkish people can be grouped in 4 classes according to 2008 IT data. These classes are: "*High IT competence*" (Class 1), "*low IT competence*" (Class 2), "*highest IT competence*" (Class 3), and "*low IT competence and highly affected by popular culture*" (Class 4). According to the results, Turkish people are most likely to be in "*low IT competence*" class (Class 2) with 0.6159 probability and least likely to be in "*highest IT competence*" (Class 3) with 0.1039 probability. This shows that population with low IT competence is dense. In order to reduce the digital divide among individuals, it necessary to increase the levels of internet and computer use. It is known that there is relationship between a country's socio-economic development and its IT level. Therefore; socio-economic developments in less developed parts of Turkey will increase the IT levels of these regions. Studies to this day showed that demographic factors such as age, gender and education affect domestic digital divide. Consequently, in order to reduce the domestic digital divide, it is necessary to implement state policies to encourage older groups to use internet and computers, to organize education programs to increase women's IT levels, and to provide vocational training for people with low levels of education.

References

[1] F. Rodrigez, J.W. Ernest, Are Poor Countries Losing the Information Revolution? *World Bank*, The Information for Development Program Working Paper, 2000.

[2] N.G. Dragulanescu, Social Impact of the "Digital Divide" in a Central-Eastern European Country. *International Information & Library Review*, 34, 139-151 (2002).

[3] D. Tapscott, *Growing up Digital: The Rise of the Net Generation*, New York: Mc - Graw Hill, 1998.

[4] R. Cullen, Adressing The Digital Divide, *Online Information Review*, 25, 5, 311-320, (2001).

[5] OECD, "Understanding Digital Divide", OECD (Organisation for Economic Co-Operation and Development), Paris, 2001.

[6] N. Corrochner, A. Ordaning, Measuring the Digital Divide: A Framework for the analysis of Cross-Country Differences. *Journal of Information Technology*, 17, 9-19, 2002.

[7] L. Öztürk, Türkiye'de Dijital Eşitsizlik: Tübitak-Bilten Anketleri Üzerine Bir Değerlendirme. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24, 111-131, 2005.

[8] M. Vicente, A. Lo´pez, Assessing the Regional Digital Divide across the European Union-27. *Telecommunications Policy*, 35, 220–237, (2011).

[9] J.B. Horrigan, C. Stolp, R.H. Wilson, Broadband Utilization in Space: Effects of Population and Economic Structure. *The Information Society*, 22, 341–356, 2006.

[10] R.W. Fairlie, D.O. Beltran, K.K. Das, Home Computers and Educational Outcomes: Evidence from The NLSY97 and CPS. *Economy Inquiry*, 48, 771–792, (2010).

[11] C. Forman, A. Goldfarb, S. Greenstein, How Did Location Affect Adoption of the Commercial Internet? Global Village vs. Urban Leadership. *Journal of Urban Economics*, 58, 389–420, (2005).

[12] Devlet Planlama Teşkilatı, "Bilgi Toplumu İstatistikleri", 2011 T.R. Ministry of Development, Statistics of Information Society, 2011.

[13] Turkish Statistical Institute, "IT Usage of Households Survey 2010", 2011.

[14] R. Özcivelek, M. Aşkar, H. Geray, T. Tüfekçi, H. Zontul, A. Küçükçınar, "Sayısal Uçurum: Dünya ve Türkiye'de Durum". *VI. Türkiye'de Internet Konferansı*, 9-11 Kasım 2000, İstanbul, 2000.

[15] E. Oruç, S. Arslan, "Sayısal Uçurumun Önlenmesi: Stratejik Plan", Telekomünikasyon Kurumu, Sektörel Araştırma ve Stratejiler Dairesi Başkanlığı, http://www.tk.gov.tr/kutuphane_ve_veribankasi/raporlar/arastirma_raporlari/dosyalar/Sayisal_Ucurumun_onlenmesi.pdf, 2002, 22 Aralık 2012.

[16] P. Aşkar, "Eğitimde Sayısal Uçurum (Digital Divide in Education)", *III.Uluslararası Eğitim Teknolojileri Sempozyumu*, Doğu Akdeniz Üniversitesi, 28-30 Mayıs, 2003.

[17] Ö. Uçkan, "Dijital Bölünme Ve Bilgi Uçurumu", *Ulusal e-devlet Konferansı*, 4-5 Kasım, Ankara, 2008.

[18] S.S. Seferoğlu, Ü. Avcı, E. Kalaycı, "Sayısal Uçurum: Türkiye'deki Durum ve Mücadelede Uygulanabilecek Politikalar", *25. Ulusal Bilişim Kurultayı*, 19-21 Kasım, Ankara, 2008.

[19] P.F. Lazarfeld, N.W. Henry, *Latent Structure Analysis*, Boston: Houghton Mifflin, 1968.

[20] N.W. Henry, "Latent Structure Analysis" in S. Kotz and N.L. Johnson (eds.) *Encyclopaedia of Statistical Sciences*, New York: Wiley, 1983, 497-504.

[21] B.F. Green, A General Solution for the Latent Class Model of Latent Structure Analysis. *Psychometrika*, 16, 151-166 (1951).

[22] B.F. Green, Latent Structure Analysis and Its Relation to Factor Analysis. *Journal of the American Statistical Association*, 47, 71-76 (1952).

[23] A.L. McCutcheoun, *Latent Class Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, No: 07-064, Newbury Park, CA: Sage, 1987.

[24] A. Agresti*,* *Categorical Data Analysis*, Wiley Series in Probability and Mathematical Statistics, Applied Probability and Statistics, John Wiley&Sons, USA, 1990.

[25] L.A. Goodman, the Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable. Part I-A: Modified Latent Structure Approach. *American Journal of Sociology*, 79, 1179-1259 (1974).

[26] L.A. Goodman, on Estimation of Parameters in Latent Structure Analysis. *Psychometrika,* 44, 123-128 (1979).

[27] T.W. Anderson, on Estimation of Parameters in Latent Structure Analysis. Psychometrika, 19, 1-10 (1954).